Appendix H Data Quality Evaluation

Table of Contents

1.0 Introduction

This appendix presents information for assembling the analytical data available after a monitoring investigation has been completed and deciding which of the data are of sufficient quality to be used in the risk assessment. Each sample may have been analyzed for the presence of many different air toxics, and many of those substances may have been detected. The following nine steps describe an approach to organize the data for use in a risk assessment. This stepwise approach is modified from that described in Chapter 5 of EPA's *Risk Assessment Guidance for Superfund*.[1] Note that the application of this stepwise approach requires considerable knowledge related to sampling and analysis methods and risk assessment and therefore should be done in consultation with appropriate experts.

1. Gather all data available from the sampling investigation and sort by medium (Section 2);
2. Evaluate the analytical methods used (Section 3);
3. Evaluate the quality of data with respect to sample quantitation limits (Section 4);
4. Evaluate the quality of data with respect to qualifiers and codes (Section 5);
5. Evaluate the quality of data with respect to blanks (Section 6);
6. Evaluate tentatively identified compounds (Section 7);
7. Compare potential contamination with background (Section 8);
8. Develop a set of data for use in the risk assessment (Section 9); and
9. Further limit the number of chemicals to be carried through the risk assessment, if appropriate (Section 10).
10. Summarize and present data (Section 11).

Acronyms for Appendix H

CLP	=	Contract Laboratory Program
CRDL	=	Contract-Required Detection Limit
CRQL	=	Contract-Required Quantitation Limit
EQL	=	Estimated Quantitation Limits
DL	=	Detection Limit
FIT	=	Field Investigation Team
IDL	=	Instrument Detection Limit
MDL	=	Method Detection Limit
ND	=	Non-detect
PE	=	Performance Evaluation
PQL	=	Practical Quantitation Limit
QA/QC	=	Quality Assurance/Quality Control
QL	=	Quantitation Limit
RfC	=	Inhalation Reference Concentration
RfD	=	Oral Reference Dose
SQL	=	Sample Quantitation Limit
SVOC	=	Semivolatile Organic Chemical
TCL	=	Target Compound List
TIC	=	Tentatively Identified Compound
TOC	=	Total Organic Carbon
TOX	=	Total Organic Halogens
VOC	=	Volatile Organic Chemical

The outcome of this evaluation is (1) the identification of contaminants of potential concern (COPC) that will be carried through the risk assessment and (2) reported concentrations that are of acceptable quality for use in a quantitative risk assessment. If the nine data evaluation steps are followed, the number of air toxics to be considered in the remainder of the risk assessment usually will be less than the number of substances initially identified. A suggested process for averaging acceptable data to develop chemical specific exposure concentrations is provided in Appendix I.

Definitions for Appendix H

Chemicals of Potential Concern. Air toxics that are evaluated in the risk assessment because they have the potential to affect the risk management decision. The corresponding term for ecological risk assessment are chemicals of potential ecological concern (COPEC). The risk assessment often finds that most of the risk is associated with a subset of the COPC. The subset, which drives the risk management decisions, is referred to as chemicals of concern (COC).

Common Laboratory Contaminants. Certain organic chemicals (e.g., acetone, 2-butanone, methylene chloride, toluene, and the phthalate esters) that are commonly used in the laboratory and thus may be introduced into a sample from laboratory cross-contamination.

Contract-required Quantitation Limit (CRQL). Chemical-specific levels that the laboratory must be able to routinely and reliably detect and quantitate in specified sample matrices to meet pre-specified data quality objectives. May or may not be equal to the reported quantitation limit of a given chemical in a given sample. (This term is also used in the Superfund Program under their Contract Laboratory Program.)

Detection Limit (DL). The lowest amount that can be distinguished from the normal "noise" of an analytical instrument or method.

Non-detects (NDs). Chemicals that are not detected in a particular sample above a certain limit, usually the quantitation limit for the chemical in that sample. Non-detects are often indicated by a "U" data qualifier.

Positive Data. Analytical results for which measurable concentrations (i.e., above a quantitation limit) are reported. May have data qualifiers attached (except a U, which indicates a non-detect).

Quantitation Limit (QL). The lowest level at which a chemical can be accurately and reproducibly quantitated. Usually equal to the instrument detection limit multiplied by a factor of three to five, but varies for different chemicals and different samples.

2.0 Step 1: Gather All Data Available from the Sampling Investigation and Sort by Medium

Gather data, which may be from several different sampling periods and based on several different analytical methods, from all available sources. Sort data by medium (i.e., air, water, sediment, soil, and biota, if appropriate). Exhibit 1 illustrates a useful table format for presenting data.

The data should be given to the risk assessor in a data summary report (or reports) that provides information on a number of critical elements that allow the assessor to judge the adequacy of the data to perform the risk analysis. Some of the critical elements include:

- Description of the study area,
- Sampling design and sampling locations,
- Procedures followed to ensure quality data (e.g., SOPs, QAPPs),
- Analytical methods and quantitation limits,

- Chemical-specific results on a per sample basis,

Exhibit 1. Example of Output Format for Validated Data

Hypothetical Soil Sampling Results from Area X

Sample medium	Soil			Soil			Soil		
Sample ID	SRB-3-1			SRB-3-1DU			SRB-3-2		
Sample or screen depth	0-1'			0-1'			2-4'		
Date collected	12/14/99			12/14/99			12/14/99		
Air Toxic	SQL[(a)]	Concen-tration	Quali-fier[(b)]	SQL[(a)]	Concen-tration	Quali-fier[(b)]	SQL[(a)]	Concen-tration	Quali-fier[(b)]
toxaphene	80	80	U	80	80	U	80	40	J
2,4,7,8-TCDD	20	10	J	20	8	J	200	200	U/J
lead	160	120	J	160	110	J	400	360	J
mercury	60	30	J	60	44	J	300	300	U/J

Note: All values other than qualifiers must be entered as numbers, not labels.
[(a)] Sample quantitation limit. Values for illustration only.
[(b)] Refer to Section 5.1 (Exhibit 3) for an explanation of qualifiers.

- Field conditions, including meteorological conditions,
- Data validation reports (both by the laboratory and any secondary validation), and
- A description of any issues with field collection, transportation/storage, or analysis that impact the veracity of the data.

The data reports provided to the risk assessor must be sufficient to allow the assessor to judge the completeness, comparability, representativeness, precision, and accuracy of the data.

[A more thorough overview of the process for assessing the usability of data for risk assessment purposes, including minimum data and documentation needs, is provided in reference 2. While this document was developed for the Superfund program, it provides relevant information for the evaluation of environmental monitoring data in a risk assessment context and, as such, is referenced here. Assessors are strongly encouraged to review this document prior to planning and scoping a assessment. This will help to ensure that all the information necessary to assess the useability of data for risk assessment purposes will be developed during the sampling and analysis phase of the assessment. (For example, assessing precision of sampling results is usually performed by establishing duplicate monitors at one or more sampling stations. The requirements for duplicate sampling must be written into the analytical plan during the planning and scoping phase of the assessment.) Reference 2 may also be consulted for information on assessing the useability of historical data for risk assessment.][(2)]

Evaluate data from different time periods to determine if concentrations of air toxics are similar or if changes have occurred between sampling periods (e.g., during different seasons of the year). If the methods used to analyze samples from different time periods are similar in terms of the types of analyses conducted and the QA/QC procedures followed, then the data may be combined for the purposes of quantitative risk assessment. Usually, this means averaging at least one year's worth of data to develop an estimate of long term average concentration (see Appendix I for a suggested methodology for combining results from air monitoring to estimate exposure concentration for the inhalation pathway). If concentrations of air toxics change significantly between sampling periods, it may be useful to also note temporal variation in the risk characterization. If data are available that spans long periods of time (e.g., multiple years) one could use only the most recent data in the quantitative risk assessment and evaluate older data in a qualitative analysis of changes in concentrations over time. When data are eliminated from a data set, justification for such elimination should be fully described in the risk assessment report. (A good understanding of the risk management goals will help in deciding what data to keep and how to combine data.)

3.0 Step 2: Evaluate the Analytical Methods Used

Group data according to the types of analyses conducted (e.g., Toxic Organic method, semivolatiles analyzed by EPA methods for air) to determine which analytical method results are appropriate for use in quantitative risk assessment.

Some types of data usually are ***not*** appropriate for use in quantitative risk assessment, even though they may be available. For example, analytical results that are not specific for a particular compound (e.g., total organic carbon [TOC], total organic halogens [TOX]), or results from insensitive analytical methods (e.g., analyses using portable field instruments such as organic vapor analyzers and other field screening methods) may be useful for identifying potential monitoring locations and/or examining the potential fate and transport of contaminants. These types of analytical results, however, generally are not appropriate for quantitative risk assessment. In addition, the results of analytical methods associated with unknown, few, or no QA/QC procedures are generally eliminated from further quantitative use. (Note that one of the purposes of the data quality objectives (DQO) process described in Chapter 6 and elsewhere in this manual is to avoid the use of sampling and analysis protocols that will not provide data that are useable for the risk assessment). These types of results, however, may be useful for qualitative discussions of risk.

The outcome of this step is a set of study-specific data that has been developed according to a standard set of sensitive, chemical-specific methods (see Chapters 10 and 19 for links to identified, standardized methods).

Note however that even when standardized, verified field and analytical procedures and associated QA/QC have been used during sampling and analysis, there is no guarantee that all analytical results are consistently of sufficient quality and reliability for use in quantitative risk assessment. Instead, it is important to determine – according to the steps discussed below – the limitations and uncertainties associated with the data, so that only data that are appropriate and reliable for use in a quantitative risk assessment are carried through the process.

4.0 Step 3: Evaluate the Quality of Data with Respect to Sample Quantitation Limits

This step involves evaluation of quantitation limits (QLs) and detection limits (DLs) for all of the air toxics assessed. This evaluation may lead to the re-analysis of some samples, the use of "proxy" (or estimated) concentrations, and/or the elimination of certain air toxics from further consideration (because they are believed to be absent in all samples). Types and definitions of QLs and DLs are presented in the box on the next page. Before eliminating an air toxic because they are not detected (or conducting any other manipulation of the data), the following points should be considered:

- The sample quantitation limit (SQL) for a specific air toxic may be greater than corresponding standards, criteria, or concentrations against which the concentrations will be compared (e.g., RfCs, RfDs, or ecological benchmark levels). In this situation, the "undetected" air toxic may be present at levels greater than these benchmarks and their exclusion from the risk assessment may result in an underestimate of risk.

- A particular SQL may be significantly higher than positively detected values in other samples in a data set.

These two points are discussed in detail in the following two subsections. A third subsection provides guidance for situations where only some of the samples for a given medium test positive for a particular chemical. A fourth subsection addresses the special situation where SQLs are not available. The final subsection addresses the specific steps involved with elimination of air toxics from the quantitative risk assessment based on their QLs.

4.1 Sample Quantitation Limits (SQLs) That Are Greater Than Benchmark Concentrations

QLs needed for the sampling and analysis investigation should be specified in the sampling plan. For some air toxics, however, SQLs obtained from available analytical methods may exceed certain concentrations of potential concern (e.g., RfCs, tissue sample concentrations that might result in a dietary intake level that exceeds an RfD). Exhibits 10-10 and 10-11 identify some known deficiencies in available air monitoring methods and some air toxics for which improved monitoring methods are needed. Two points should be noted when considering this situation:

- Review of available information on sources and emissions, a preliminary determination of COPC, and/or the results of fate and transport modeling ***prior to sample collection*** may allow the risk assessor to identify when more sensitive sampling and/or analytical methods may be needed before an investigation begins. This is the most efficient way to minimize the problem of QLs exceeding levels of potential concern.

- Analytical laboratories may not be able to attain QLs in particular samples that meet data quality requirements using standardized, verified procedures.

If an air toxic is not detected in any sample from a particular medium at the QL and a more sensitive method is not available, then modeling data, as well as professional judgment, may be used to evaluate whether the chemical may be present above the concentrations of potential concern. If the available information indicates the chemical is not present, see Section 3.5 of this

Detection Limits and Quantitation Limits

Strictly interpreted, the detection limit (DL) is the lowest amount of a chemical that can be "seen" above the normal, random noise of an analytical instrument or method. A chemical present below that level cannot reliably be distinguished from noise. DLs are chemical-specific and instrument-specific and are determined by statistical treatment of multiple analyses in which the ratio of the lowest amount observed to the electronic noise level (i.e., the signal-to-noise ratio) is determined. On any given day in any given sample, the calculated limit may not be attainable; however, a properly calculated limit can be used as an overall general measure of laboratory performance.

Two types of DLs may be described: instrument DLs (IDLs) and method DLs (MDLs). The IDL is generally the lowest amount of a substance that can be detected by an instrument; it is a measure only of the DL for the instrument, and does not consider any effects that sample matrix, handling, and preparation may have. The MDL, on the other hand, takes into account the reagents, sample matrix, and preparation steps applied to a sample in specific analytical methods.

Due to the irregular nature of instrument or method noise, reproducible quantitation of a chemical is not possible at the DL. Generally, a factor of three to five is applied to the DL to obtain a quantitation limit (QL), which is considered to be the lowest level at which a chemical may be accurately and reproducibly quantitated. DLs indicate the level at which a small amount would be "seen," whereas QLs indicate the levels at which measurements of concentration can be "trusted."

Two types of QLs may be described: estimated quantitation limits (EQL - also sometimes referred to as a practical quantitation limit or PQL) and sample QLs (SQLs). EPA's Superfund Program maintains a Contract Laboratory Program (CLP) as a means to obtain reliable analytical results from many different laboratories. To participate in the CLP, a laboratory must be able to meet EPA's EQL. This EQL is established by contract and, thus, is called a contract required quantitation limit (CRQL). CRQLs are chemical-specific and vary depending on the medium analyzed and the amount of chemical expected to be present in the sample. As the name implies, CRQLs are not necessarily the lowest detectable levels achievable, but rather are levels that a CLP laboratory should routinely and reliably detect and quantitate in a variety of sample matrices. For most air toxics risk assessments, SQLs, not CRQLs, will be the QLs of interest for most samples. In fact, for the same chemical, a specific SQL may be higher than, lower than, or equal to SQL values for other samples. In addition, preparation or analytical adjustments such as dilution of a sample for quantitation of an extremely high level of only one compound could result in non-detects for all other compounds included as analytes for a particular method, even though these compounds may have been present at trace quantities in the environmental sample. Because SQLs take into account sample characteristics, sample preparation, and analytical adjustments, these values are the most relevant QLs for evaluating non-detected chemicals. Also note that because of the inability to accurately measure concentration at the MDL, the SQL is used as he starting point for developing exposure concentrations where some of the samples in a data set have detections of an analyte and others do not (see Appendix I).

appendix for guidance on eliminating chemicals. If there is some indication that the chemical is present, the only choices are to:

- Use modeling results in the risk assessment;
- Re-analyze selected samples using a more sensitive analytical method (if feasible); or
- Address the chemical qualitatively in the risk assessment.

In determining which option is most appropriate for an analysis, it may be helpful to assume the air toxic is present at the SQL for purposes of an initial (tier 1) screening risk assessment. In this way, risks that would be posed if the chemical is present at the SQL can be compared with risks posed by other air toxics in the analysis.

4.2 Unusually High SQLs

Due to one or more sample-specific problems (e.g., matrix interferences), SQLs for a particular chemical in some samples may be unusually high, sometimes greatly exceeding the positive results reported for the same chemical in other samples from the data set. Even if these SQLs do not exceed health-based standards or criteria, they may still present problems. If the SQLs cannot be reduced by re-analyzing the sample, consider excluding the samples from the quantitative risk assessment if they cause the calculated exposure concentration to exceed the maximum detected concentration for a particular sample set. Exhibit 2 presents an example of how to address a situation with unusually high QLs.

Exhibit 2. Example of Unusually High Quantitation Limits

In this hypothetical example, ambient air concentrations of benzene in air have been determined using the TO-1 method.

	Concentration (ppb)			
Chemical	**Sample 1**	**Sample 2**	**Sample 3**	**Sample 4**
benzene	50 U[(a)]	59	200 U	74

[(a)] U indicates that benzene was analyzed for, but not detected; the value presented (e.g., 50 U) is the SQL.

The ambient air concentrations presented in this example (i.e., 50 to 200 ppb) vary widely from sample to sample. Assume a more sensitive analytical method would not aid in reducing the unusually high QL of 200 ppb noted in Sample 3. In this case, the result for benzene in Sample 3 would be eliminated from the quantitative risk assessment because it would cause the calculated exposure concentrations to exceed the maximum detected concentration (in this case 74 ppb). Thus the data set would be reduced to three samples: the non-detect in Sample 1 and the two detected values in Samples 2 and 4.

4.3 When Only Some Samples in a Medium Test Positive For a Chemical

Most analytes are not positively detected in each sample collected and analyzed. Instead, for a particular chemical the data set generally will contain some samples with positive results and others with non-detected results. The non-detected results usually are reported as SQLs. These limits indicate that the chemical was not measured above certain levels, which may vary from sample to sample. The chemical may be present at a concentration just below the reported quantitation limit, or it may not be present in the sample at all (i.e., the concentration in the sample is zero). Appendix I provides a suggested methodology for combining the results of a dataset where some of the samples test positive for an analyte and others do not.

4.4 When SQLs Are Not Available

In some cases, laboratory data summaries may not provide the SQLs. Instead, MDLs, CRQLs, or even IDLs may have been substituted wherever a chemical was not detected. Sometimes, no detection or quantitation limits may be provided with the data. As a first step in these situations, always attempt to obtain the SQLs, because these are the most appropriate limits to consider when evaluating non-detected air toxics (i.e., they account for sample characteristics, sample preparation, or analytical adjustments that may differ from sample to sample). Good planning and clearly articulated directions to the laboratory will help ensure that the appropriate information is provided to the risk assessor. The problem associated with incorrectly reported data should only be an issue when evaluating historical data for which there was no pre-consultation with the laboratory about what is to be provided in the data package.

If SQLs cannot be obtained, the MDL may be used as the QL, with the understanding that in most cases this will underestimate the SQL (because the MDL is a measure of detection limits only and does not account for sample characteristics or matrix interferences). The IDL should rarely be used for non-detected air toxics since it is a measure only of the detection limit for a particular instrument and does not consider the effect of sample handling and preparation or sample characteristics.

4.5 When Air Toxics Are Not Detected in Any Samples in a Medium

After considering the discussion provided in the above subsections, generally eliminate those air toxics that have not been detected in any samples of a particular medium. If information exists to indicate that the air toxics are present, they should not be eliminated from the analysis. The outcome of this step is a data set that only contains air toxics for which positive data (i.e., analytical results for which measurable concentrations are reported) are available in at least one sample from each medium. Unless otherwise indicated, assume at this point in the evaluation of data that positive data to which no uncertainties are attached concerning either the assigned identity of the chemical or the reported concentration (i.e., data that are not “tentative,” “uncertain,” or “qualitative”) are appropriate for use in the quantitative risk assessment.

5.0 Step 4: Evaluate the Quality of Data with Respect to Qualifiers and Codes

Various qualifiers and codes (hereafter referred to as qualifiers) may be attached to certain data by either the laboratories conducting the analyses or by persons performing data validation. These qualifiers often pertain to QA/QC problems and generally indicate questions concerning

chemical identity, chemical concentration, or both. All qualifiers must be addressed before the chemical can be used in quantitative risk assessment. Qualifiers used by the laboratory may differ from those used by data validation personnel in either identity or meaning.

5.1 Types of Qualifiers

Exhibit 3 provides a list of the qualifiers that laboratories are permitted to use under the Superfund CLP, along with their potential use in risk assessment. Exhibit 4 provides a similar list addressing data validation qualifiers. (Note that the data qualifiers and their meanings provided here are not consistent across all laboratories. In all cases, it is critical to discuss with the lab what they mean by the data qualifiers they report.) In general, because the data validation process is intended to assess the effect of QC issues on data usability, validation data qualifiers are attached to the data after the laboratory qualifiers and supersede the laboratory qualifiers. If data have both laboratory and validation qualifiers and they appear contradictory, ignore the laboratory qualifier and consider only the validation qualifier. If qualifiers have been attached to certain data by the laboratory and have not been removed, revised, or superseded during data validation, then evaluate the laboratory qualifier itself. If it is unclear whether the data have been validated, contact the appropriate data validation and/or laboratory personnel.

The type of qualifier and other site-specific factors determine how qualified data are to be used in a risk assessment. As seen in Exhibits 3 and 4, the type of qualifier attached to certain data often indicates how that data should be used in a risk assessment. For example, most of the laboratory qualifiers for both inorganic chemical data and organic chemical data (e.g., J, E, N) indicate uncertainty in the reported concentration of the chemical, but not in its assigned identity. Therefore, these data can be used just as positive data with no qualifiers or codes. In general, include data with qualifiers that indicate uncertainties in concentrations but not in identification.

Exhibit 3. Example of Data Qualifiers and Their Potential Use in Quantitative Risk Assessment: Superfund Contract Laboratory Program (CLP)

Qualifier	Definition	Indicates:		
		Uncertain Identity?	Uncertain Concentration?	Include Data in Quantitative Risk Assessment?
Inorganic Chemical Data[(a)]				
B	Reported value is <CRDL, but >IDL.	No	No	Yes
U	Compound was analyzed for, but not detected.	Yes	Yes	?
E	Value is estimated due to matrix interferences.	No	Yes	Yes
M	Duplicate injection precision criteria not met.	No	Yes	Yes

Exhibit 3. Example of Data Qualifiers and Their Potential Use in Quantitative Risk Assessment: Superfund Contract Laboratory Program (CLP)				
Qualifier	Definition	Indicates:		
		Uncertain Identity?	Uncertain Concentration?	Include Data in Quantitative Risk Assessment?
N	Spiked sample recovery not within control limits.	No	Yes	Yes
S	Reported value was determined by the Method of Standard Additions (MSA).	No	No	Yes
W	Post-digestion spike for furnace AA analysis is out of control limits, while sample absorbance is <50% of spike absorbance.	No	Yes	Yes
*	Duplicate analysis was not within control limits.	No	Yes	Yes
+	Correlation coefficient for MSA was <0.995.	No	Yes	Yes
Organic Chemical Data[(b)]				
U	Compound was analyzed for, but not detected.	Yes	Yes	?
J	Value is estimated, either for a tentatively identified compound (TIC) or when a compound is present (spectral identification criteria are met, but the value is <CRQL).	No for TCL chemicals Yes for TICs	Yes	?
C	Pesticide results were confirmed by GC/MS.	No	No	Yes
B	Analyte found in associated blank as well as in sample.[©)]	No	Yes	Yes
E	Concentration exceeds calibration range of GC/MS instrument.	No	Yes	Yes
D	Compound identified in an analysis at a secondary dilution factor.	No	No	Yes
A	The TIC is a suspected aldol-condensation product.	Yes	Yes	No

Exhibit 3. Example of Data Qualifiers and Their Potential Use in Quantitative Risk Assessment: Superfund Contract Laboratory Program (CLP)				
Qualifier	Definition	Indicates:		
		Uncertain Identity?	Uncertain Concentration?	Include Data in Quantitative Risk Assessment?
X	Additional flags defined separately.	–(d)	--	--

(a) Source: U.S. EPA, 1988. *Contract Laboratory Program Statement of Work for Inorganics Analysis: Multi-media, Multi-concentration.* Office of Emergency and Remedial Response. SOW No. 788.
(b) Source: U.S. EPA, 1988. *Contract Laboratory Program Statement of Work for Organics Analysis: Multi-media, Multi-concentration.* Office of Emergency and Remedial Response. SOW No. 288.
©) See Section 6 for a discussion of blank contamination.
(d) Data will vary with laboratory conducting analyses.

Exhibit 4. Validation Data Qualifiers and Their Potential Use in Quantitative Risk Assessment				
Qualifier	Definition	Indicates:		
		Uncertain Identity?	Uncertain Concentration?	Include Data in Quantitative Risk Assessment?
Inorganic and Organic Chemical Data[(a)]				
U	The material was analyzed for, but not detected. The associated numerical value is the SQL.	Yes	Yes	?
J	The associated numerical value is an estimated quantity.	No	Yes	Yes
R	Quality control indicates that the data are unusable (compound may or may not be present). Re-sampling and/or re-analysis is necessary for verification.	Yes	Yes	No
Z	No analytical result (inorganic data only).	--	--	--
Q	No analytical result (organic data only).	--	--	--
N	Presumptive evidence of presence of material (tentative identification)[(b)]	Yes	Yes	?

Exhibit 4. Validation Data Qualifiers and Their Potential Use in Quantitative Risk Assessment

(a) Source: U.S. EPA. 1988. *Laboratory Data Validation Functional Guidelines for Evaluating Inorganics Analysis*. Office of Emergency and Remedial Response.
U.S. EPA. 1988. *Laboratory Data Validation Functional Guidelines for Evaluating Organics Analysis (Functional Guidelines for Organics).* Office of Emergency and Remedial Response.
(b) Organic chemical data only

Exhibit 5 provides examples showing the use of two commonly encountered data qualifiers: the J qualifier, and the R qualifier. Basically, the suggestion is to use J-qualified concentrations the same way as positive data that do not have this qualifier. If possible, note potential uncertainties associated with the qualifier, so that if data qualified with a J contribute significantly to the risk, then appropriate caveats can be attached. The R data qualifier indicates that the sample result was rejected by the data validation personnel, and therefore this result should be eliminated from the risk assessment.

Exhibit 5. Example Use of "J" and "R" Data Qualifiers

In this example, concentrations of benzene in an air monitor have been determined using a hypothetical analytical method. Benzene was detected in these four samples at concentrations of 3,200 µg/l, 40 µg/l, and 20 µg/l; therefore, these concentrations – as well as the non-detect – should be used in determining representative concentrations.

Chemical	Sample 1	Sample 2	Sample 3	Sample 4
Benzene	3,200 J[(a)]	40	30 U[(b)]	20 J

(a) J = The numerical value is an estimated quantity
(b) U = Compound was analyzed for, but not detected. Value presented (e.g., 30 U) is the SQL.

In this example, concentrations of lead in surface water have been determined using a hypothetical analytical method. These data have been validated, and therefore the R qualifers indicate that the person conducting the data validation rejected the data for lead in samples 2 and 3. The "UR" qualifier means that lead was not detected in Sample 3; however, the data validator rejected the non-detected result. Eliminate these two samples so that the data set now consists of only two samples (Samples 1 and 4).

Chemical	Sample 1	Sample 2	Sample 3	Sample 4
Lead	310	500 R[(a)]	30 UR[(b)]	500

(a) R = Quality control indicates that the data are unusable (compound may not be present)
(b) U = Compound was analyzed for, but not detected. Value presented (e.g., 30 UR) is the SQL.

5.2 Using the Appropriate Qualifiers

The information presented in Exhibits 3 and 4 is based on 1988 EPA guidance documents concerning qualifiers. The types and definitions of qualifiers may be periodically updated within any analytical program, and EPA regions, states, and local governments may have their own data

qualifiers and associated definitions. In general, the risk assessor should clearly understand the specific data qualifiers used by a particular analytical program and use the resulting data appropriately in the risk assessment. Make sure that definitions of data qualifiers used in the data set for the analysis have been reported with the data and are current. Never guess about the definition of qualifiers.

6.0 Step 5: Evaluate the Quality of Data with Respect to Blanks

Blank samples provide a measure of contamination that has been introduced into a sample set either (1) in the field while the samples were being collected or transported to the laboratory, or (2) in the laboratory during sample preparation or analysis. To prevent the inclusion of non-site-related contaminants in the risk assessment, the concentrations of air toxics detected in blanks must be compared with concentrations of the same air toxics detected in site samples. Exhibit 6 provides detailed definitions of different types of blanks. Blank data should be compared with results from samples with which the blanks are associated. It is often impossible, however, to determine the association between certain blanks and data. In this case, compare the blank data with results from the entire sample data set. EPA's Superfund Program has developed guidelines for comparing sample concentrations with blank concentrations; **note that the requirements or practices for a given air toxic program may differ**.

- **Blanks containing common laboratory contaminants**. As discussed in the EPA documents cited in Exhibits 3 and 4, acetone, 2- butanone (or methyl ethyl ketone), methylene chloride, toluene, and the phthalate esters are considered by EPA to be common laboratory contaminants. If the blank contains detectable levels of common laboratory contaminants, EPA guidance indicates that the sample results should be considered as positive results only if the concentrations in the sample exceed **ten times** the maximum amount detected in any blank. If the concentration of a common laboratory contaminant is less than ten times the blank concentration, then EPA guidance indicates to conclude that the chemical was not detected in the particular sample and consider the blank-related concentrations of the chemical to be the quantitation limit for the chemical in that sample. Note that if all samples contain levels of a common laboratory contaminant that are less than ten times the level of contamination noted in the blank, then completely eliminate that chemical from the set of sample results.

- **Blanks containing chemicals that are not common laboratory contaminants**. As discussed in the previously referenced guidance, if the blank contains detectable levels of one or more organic or inorganic chemicals that are not considered by EPA to be common laboratory contaminants, then consider sample results as positive only if the concentration of the chemical in the sample exceeds **five times** the maximum amount detected in any blank. Treat samples containing less than five times the amount in any blank as non-detects, and consider the blank-related chemical concentration to be the quantitation limit for the chemical in that sample. Again, note that if all samples contain levels of a chemical that are less than five times the level of contamination noted in the blank, then completely eliminate that chemical from the set of sample results.

Exhibit 6. Types of Blanks

Blanks are analytical quality control samples analyzed in the same manner as site samples. They are used in the measurement of contamination that has been introduced into a sample either (1) in the field while the samples were being collected or transported to the laboratory or (2) in the laboratory during sample preparation or analysis. Four types of blanks – trip, field, laboratory calibration, and laboratory reagent (or method) – are described below. A discussion on the water used for the blank also is provided.

Trip Blank. This type of blank is used to indicate potential contamination due to migration of volatile organic chemicals (VOCs) from the air on the site or in sample shipping containers, through the septum or around the lid of sampling vials, and into the sample. A trip blank consists of laboratory distilled, deionized water in a 40-ml glass vial sealed with a teflon septum. The blank accompanies the empty sample bottles to the field as well as the samples returning to the laboratory for analysis; it is not opened until it is analyzed in the lab with the actual site samples. The containers and labels for trip blanks should be the same as the containers and labels for actual samples, thus making the laboratory "blind" to the identity of the blanks.

Field Blank. A field blank is used to determine if certain field sampling or cleaning procedures (e.g., insufficient cleaning of sampling equipment) result in cross-contamination of site samples. Like the trip blank, the field blank is a sample of distilled, deionized water taken to the field with empty sample bottles and is analyzed in the laboratory along with the actual samples. Unlike the trip blank, however, the field blank sample is opened in the field and used as a sample would be (e.g., it is poured through cleaned sampling equipment or it is poured from container to container in the vicinity of a gas-powered pump). As with trip blanks, the field blanks' containers and labels should be the same as for actual samples.

Laboratory Calibration Blank. This type of blank is distilled, deionized water injected directly into an instrument without having been treated with reagents appropriate to the analytical method used to analyze actual site samples. This type of blank is used to indicate contamination in the instrument itself, or possibly in the distilled, deionized water.

Laboratory Reagent or Method Blank. This blank results from the treatment of distilled, deionized water with all of the reagents and manipulations (e.g., digestions or extractions) to which site samples will be subjected. Positive results in the reagent blank may indicate either contamination of the chemical reagents or the glassware and implements used to store or prepare the sample and resulting solutions. Although a laboratory following good laboratory practices will have its analytical processes under control, in some instances method blank contamination cannot be entirely eliminated.

Water Used for Blanks. For all the blanks described above, results are reliable only if the water comprising the blank was clean. For example, if the laboratory water comprising the trip blank was contaminated with VOCs prior to being taken to the field, then the source of VOC contamination in the trip blank cannot be isolated (see laboratory calibration blank).

7.0 Step 6: Evaluate Tentatively Identified Compounds

Both the identity and reported concentration of a tentatively identified compound (TIC) is questionable (see Exhibit 7). Two options for addressing TICs exist, depending on the relative number of TICs compared to non-TICs. If the risk assessment involves a regulatory decision, the risk assessor is strongly encouraged to consult the appropriate regulatory authorities about how to address TICs in the risk assessment.

- **When few TICs are present**. When only a few TICs are present, and either (a) no information indicates that either a particular TIC may indeed be present (e.g., it is not present in emissions from the source(s) being evaluated or other nearby sources), or (b) the estimated concentration is relatively low, and therefore, the risk estimate would likely not be dominated by the TIC, then generally do not include the TICs in the risk assessment.

- **When Many TICs are present**. If many TICs are present, or if TIC concentrations appear high or site information indicates that TICs are indeed present, then further evaluation of TICs is necessary. If sufficient time is available, use more sensitive analytical methods to confirm the identity and to positively and reliably measure the concentrations of TICs prior to their use in the risk assessment. If such methods are unavailable or impractical, then the TICs should be included as COPC in the risk assessment and (usually) discussed qualitatively in the risk characterization along with a discussion of the uncertainty in both identity and concentration.

Exhibit 7. Tentatively Identified Compounds (TICs)

The set of compounds analyzed in a particular laboratory protocol may be a limited subset of the organic air toxics that could actually be present in specific emissions being evaluated. Thus, a laboratory analysis may indicate the presence of additional organic compounds not being specifically evaluated. The presence of additional compounds may be indicated, for example, by"peaks" on a chromatogram (a chromatogram is a paper representation of the response of the instrument to the presence of a compound). The laboratory may be required to attempt to identify some of these compounds (e.g., the highest peaks) using computerized searches of a library containing mass spectra (essentially "fingerprints" for particular compounds). When the mass spectra match to a certain degree, the compound (or general class of compound) is named; however, the assigned identity is in most cases highly uncertain. These compounds are called tentatively identified compounds (TICs).

The analytical protocols being used by the laboratory may include procedures to obtain a rough estimate of the concentrations of TICs. These estimates, however, generally are highly uncertain and could be orders of magnitude higher or lower than the actual concentration. For TICs, therefore, assigned identities may be inaccurate, and quantitation is certainly inaccurate. Due to these uncertainties, TIC information often is not provided with data summaries. Additional sampling and analysis using different or more sensitive methods may reduce the uncertainty associated with TICs and, therefore, TIC information should be sought even if it is absent from data summaries.

8.0 Step 7: Compare Potential Contamination with Background

In some cases, a comparison of sample concentrations with background concentrations is useful for identifying the relative contribution of the source(s) being evaluated and other potential sources to the total concentrations to which a population may be exposed. Often, however, the comparison of samples with background is unnecessary because the risk estimates resulting from other sources are very low compared to those resulting from the source(s) being evaluated.

Information collected during the risk assessment can provide information on two types of background chemicals: (1) naturally occurring chemicals that have not been influenced by humans and (2) chemicals that are present due to anthropogenic sources. Either type of background chemical can be either localized or widespread. Information on background chemicals may have been obtained by the collection of background samples and/or from other

sources (e.g., County Soil Conservation Service surveys, United States Geological Survey reports). Background concentrations should be from the vicinity of the location sampled. For example, background air samples are generally collected upwind from the study area to estimate concentrations of chemicals in the air mass that is moving into the study area. For water, samples are taken upstream of the area where deposition (or erosion of contaminated soils) is occurring.

Background samples collected during the monitoring effort should not be used if they were obtained from areas influenced or potentially influenced by the source(s) being evaluated. Instead, the literature sources mentioned in the previous paragraph may be consulted to determine expected background levels of air toxics in the study area. Care must be taken in using literature sources, because the data contained therein might represent nationwide variation in a particular parameter rather than variation typical of the geographic region or geological setting in which the site is located. For example, a literature source providing concentrations of chemicals in soil on a national scale may show a wide range of concentrations that is not representative of the variation in concentrations that would be expected within a particular study area.

Both the concentration of the chemical in the study-area and the concentration in background media should be clearly articulated in the risk assessment report. Background concentrations should generally not be subtracted from study-area specific concentrations; rather, they should be compared (e.g., as barcharts). Statistical analyses that indicate whether study-area and background concentrations are different may also be presented. (In cases where background comparisons will be made, the statistical methods that will be used to compare study-area concentrations to background concentrations should be identified prior to the collection of samples.)

As an example, chromium is present in air releases from a source in a study area and chromium is also naturally occurring in study area soils. In this case, it may be necessary to include a careful comparison of the relative magnitude of estimated exposure and risk due to background vs. estimated exposure and risk from total (i.e., deposited chromium + background chromium) . This can be done by the bar chart method mentioned above and may be augmented by statistical analyses that attempt to answer the question about whether study area soil concentrations of chromium are statistically different from background soils. Again, consultation with the appropriate decision making authorities is strongly encouraged to ensure that they get the type of information that they will need to make their risk management decisions. (Note that, in general, comparison with naturally occurring levels is commonly performed primarily for inorganic chemicals such as metals, because the majority of organic air toxics released to the environment are not naturally occurring (even though they may be ubiquitous). Similar to naturally occurring background concentrations, anthropogenic levels resulting from human sources (other than those being evaluated in the air toxics risk assessment) may also be present. For example, an assessment that is evaluating exposures to dioxin from a specific source may also have to contend with dioxin that is also present in the study area that has resulted from numerous other small sources in the area (and possibly also from naturally occurring sources such as forest fires and some amount of longer range transport). Similar to naturally occurring chemicals, some combination of background sampling, literature values, modeling, and statistical analysis can be performed to try and sort out how much of the concentrations and risk are due to the source(s) in question and how much is present due to other human (and non-human) influences.

9.0 Step 8: Develop a Set of Data for Use in the Risk Assessment

After the evaluation of data is complete as specified in previous sections, a list of the samples (by medium) is made that will be used to estimate exposure concentrations. In addition, a list of COPC (also by medium) will be needed for the quantitative risk assessment. This list should include chemicals that were:

- Positively detected in at least one sample in a given medium, including (a) chemicals with no qualifiers attached (excluding samples with unusually high detection limits), and (b) chemicals with qualifiers attached that indicate known identities but unknown concentrations (e.g., J-qualified data);

- Detected at levels significantly elevated above levels of the same chemicals detected in associated blank samples;

- Only tentatively identified but either may be associated with emissions from the source(s) being evaluated based on ancillary information or have been confirmed by additional analysis; and/or

- Transformation products of air toxics demonstrated to be present.

Air toxics that were not detected in samples from a given medium (i.e., non-detects) but that may be present at the site also may be included in the risk assessment if an evaluation of the risks potentially present at the detection limit is desired.

10.0 Step 9: Further Limit the Number of Chemicals to Be Carried Through the Risk Assessment, If Appropriate

For certain assessments, the list of air toxics potentially related to emissions from the source(s) being evaluated and remaining after quantitation limits, qualifiers, blank contamination, and background have been evaluated may be lengthy. ***Note, however, that often a modeling analysis can identify the subset of air toxics in the emissions being evaluated that are most likely to contribute significantly to risk, and therefore limit the scope of any subsequent sampling and analysis effort***. Carrying a large number of chemicals through a quantitative risk assessment may be complex, and it may consume significant amounts of time and resources. The resulting risk assessment report may be difficult to read and understand, and it may distract from the dominant risks. In these cases, the procedures discussed in this section – using chemical classes, frequency of detection, essential nutrient information, and a concentration toxicity screen – may be used to further reduce the number of COPC in each medium.

If conducting a risk assessment on a large number of chemicals is feasible (e.g., because of adequate computer capability), then the procedures presented in this section may be omitted. However, the most important chemicals (e.g., those presenting 99 percent of the risk) – identified after the risk assessment – may be the focus of the main text of the report, and the remaining chemicals could be presented in the appendices.

10.1 Conduct Initial Activities

There are several activities that are useful to conduct before implementing any of the procedures described in this section. The risk assessor is strongly encouraged to consult with appropriate decision making authorities prior to implementing these procedures to ensure that the resulting processed data will meet the decision makers' needs. These remaining initial activities include:

- **Considering how the rationale for the procedure should be documented**. The rationale for eliminating chemicals from the quantitative risk assessment based on the procedures discussed below should be clearly stated in the risk assessment report. This documentation, and its possible defense at a later date, could be fairly resource- intensive. If a continuing need to justify this step is expected, then any plans to eliminate chemicals should be reconsidered.

- **Examining historical information about the source(s) being evaluated**. Chemicals reliably associated with emissions from the source(s) being evaluated based on historical information generally should not be eliminated from the quantitative risk assessment (at least during the initial tiers of analysis), even if the results of the procedures given in this section indicate that such an elimination is possible.

- **Considering mobility, persistence, and bioaccumulation**. Three factors that should be considered are the mobility, persistence, and bioaccumulation of the chemicals. For example, a highly volatile (i.e., mobile) chemical such as benzene, a long-lived (i.e., persistent) chemical such as dioxin, or a readily bioaccumulated chemical such as the PB-HAPs, probably should remain in the risk assessment. These procedures do not explicitly include a mobility, persistence, or bioaccumulation component, and therefore the risk assessor must pay special attention to these factors.

- **Considering special exposure routes**. For some chemicals, certain exposure routes need to be considered carefully before using these procedures. For example, some air toxics may pose a significant risk in certain circumstances due to dermal contact. The procedures described in this section may not account for exposure routes such as this.

10.2 Group Chemicals by Class

Some dose-response values used in characterizing risks are available only for certain chemicals within a chemical class. For example, slope factors are available only for some of the polycyclic aromatic hydrocarbons (PAHs). In such cases, the information provided in Chapter 12 (toxicity evaluation) and information provided on EPA's FERA website (http://www.epa.gov/ttn/fera/).

10.3 Evaluate Frequency of Detection

Chemicals that are infrequently detected may be artifacts in the data due to sampling, analytical, or other problems, and therefore may not be related to the sources being evaluated. Consider the chemical as a candidate for elimination from the quantitative risk assessment if: (1) it is detected infrequently in one or perhaps two environmental media, (2) it is not detected in any other sampled media or at high concentrations, and (3) there is no reason to believe that the chemical may be present in emissions from the source(s) being evaluated. In particular, modeling results may indicate whether monitoring data that show infrequently detected chemicals are representative of only their sampling locations or of broader areas. Because chemical

concentrations within a broad assessment area are spatially variable, the risk assessor can use modeling results to compare infrequently detected chemical concentrations to those estimated over broader areas when determining whether the subject chemicals are relevant to the overall risk assessment. Judicious use of modeling to supplement available monitoring data often can minimize the need to resort to arbitrarily setting limits on inclusion of infrequently detected chemicals in the risk assessment.

In addition to available monitoring data and modeling results, the risk assessor should consider other relevant factors (e.g., presence of sensitive subpopulations) in recommending appropriate site-specific limits on inclusion of risk assessment.

The reported or modeled concentrations and locations of chemicals should be examined to check for "hotspots" (localized areas of particularly high concentrations), which may be especially important for short-term exposures and which therefore should not be eliminated from the risk assessment. For PB-HAPs, always consider detection of particular chemicals in all sampled media because some media may be sources of contamination for other media. In addition, infrequently detected chemicals with concentrations that greatly exceed reference concentrations should not be eliminated.

10.4 Use a Toxicity-Weighted or Risk-based Screening Analysis

The objective of this screening procedure is to identify the chemicals in a particular analysis that, based on concentration and toxicity, are most likely to contribute significantly to the resulting risk estimates. These procedures are described, along, with examples, in Chapter 6.

11.0 Summarize and Present Data

The section of the risk assessment report summarizing the results of the data collection and evaluation should be titled "Identification of COPC." Information in this section should be presented in ways that readily support the calculation of exposure concentrations in the exposure assessment portion of the risk assessment. Exhibits 8 and 9 present examples of tables to be included in this section of the risk assessment report.

11.1 Summarize Data Collection and Evaluation Results in Text

In the introduction for this section of the risk assessment report, clearly discuss in bullet form the steps involved in data evaluation. If the optional screening procedure described in Section 9 was used in determining COPC, these steps should be included in the introduction. If both historical data and current data were used in the data evaluation, state this in the introduction. Any special site-specific considerations in collecting and evaluating the data should be mentioned. General uncertainties concerning the quality associated with either the collection or the analysis of samples should be discussed so that the potential effects of these uncertainties on later sections of the risk assessment can be determined.

In the next part of the report, discuss the samples from each medium selected for use in quantitative risk assessment. Provide information concerning the sample collection methods used (e.g., grab, composite) as well as the number and location of samples. If any samples (e.g., field screening/analytical samples) were excluded specifically from the quantitative risk

assessment prior to evaluating the data, document this along with reasons for the exclusion. Again, remember that such samples, while not used in the quantitative risk assessment, may be useful for qualitative discussions and therefore should not be entirely excluded from the risk assessment.

Discuss the data evaluation within the appropriate context for the risk assessment. For example, the focus may be on a particular neighborhood within the assessment area; specific types of modeled receptors; or specific geographic features such as a water body. For PB-HAPs, the discussion should include those media (e.g., wastes, soils) that are potential sources of contamination for other media (e.g., surface water/sediments). If no samples or data were available for a particular medium, discuss this in the text. For soils data, discuss surface soil results separately from those of subsurface soils. Discuss surface water/sediment results by the specific surface water body sampled.

Exhibit 8. Example of Table Format for Presenting Air Toxics Sampled in Specific Media

Air Toxic	Concentration in Medium X			
	Frequency of Detection[(a)]	Range of Sample Quantitation Limits (SQLs) (units)	Range of Detected Concentrations (units)	Background Levels
Chemical A	3/25	2 - 30	320 - 4600	100 - 140
Chemical B[(b)]	25/25	1 - 32	17 - 72	--

-- Not sampled
[(a)] Number of samples in which the chemical was positively detected over the number of samples available
[(b)] Identified as a COPC based upon evaluation of data according to procedures described in text of report

For each medium, identify in the report the chemicals for which samples were analyzed, and list the analytes that were detected in at least one sample. If any detected chemicals were eliminated from the quantitative risk assessment based on evaluation of data (i.e., based on evaluation of data quality, background comparisons, and the optional screening procedures, if used), provide reasons for the elimination in the text (e.g., chemical was detected in blanks at similar concentrations to those detected in samples or chemical was infrequently detected).

Exhibit 9. Example of Table Format for Summarizing COPC in All Media Sampled				
Air Toxic	Concentration			
	Air ($\mu g/m^3$)	Soils (mg/kg)	Surface Water (µg/l)	Sediments (µg/l)
Chemical A	0.5 - 225	5 - 1,100	2 - 30	--
Chemical B	0.1 - 22	0.5 - 6.4	--	12 - 3650
Chemical C	0.01 - 2.2	--	50 - 440	100 - 11,000
Chemical D	3 - 854	2 - 12	--	
-- Not sampled				

The final subsection of the text is a discussion of general trends in the data results. For example, the text may mention (1) whether concentrations of COPC in most media were close to the detection limits or (2) trends concerning chemicals detected in more than one medium or in more than one operable unit at the site. In addition, the location of hot spots should be discussed, as well as any noticeable trends apparent from sampling results at different times.

11.2 Summarize Data Collection and Evaluation Results in Tables and Graphics

As shown in Exhibit 8, a separate table that includes all chemicals detected in a medium can be provided if appropriate. Chemicals that have been determined to be of potential concern based on the data evaluation should be designated in the table with an asterisk to the left of the chemical name.

For each chemical, present the frequency of detection in a certain medium (i.e., the number of times a chemical was detected over the total number of samples considered) and the range of detected or quantified values in the samples. Do not present the QL or similar indicator of a minimum level (e.g., <10 mg/L, ND) as the lower end of the range; instead, the lower and upper bound of the range should be the minimum and maximum detected values, respectively. The range of reported QLs obtained for each chemical in various samples should be provided in a separate column. Note that these QLs should be sample-specific; other types of non-sample-specific values (e.g., MDLs or CRQLs) should be provided only when SQLs are not available. Note that the range of QLs would not include any limit values (e.g., unusually high QLs) eliminated based on the guidance in Section 3. Finally, naturally occurring concentrations of chemicals used in comparing sample concentrations may be provided in a separate column. The source of these naturally occurring levels should be provided in a footnote. List the identity of the samples used in determining concentrations presented in the table in an appropriate footnote.

The final table in this section is a list of the COPC presented by medium at the site or by medium within each operable unit at the site. A sample table format is presented in Exhibit 9. Ths isopleth is another useful type of presentation of chemical concentration data (not shown). This

graphic characterizes the monitored or modeled concentrations of chemicals at a site and illustrates the spatial pattern of contamination.

References

1. U.S. Environmental Protection Agency. 1989. *Risk Assessment Guidance for Superfund: Volume I. Human Health Evaluation Manual (Part A).* Office of Emergency and Remedial Response. Washington, DC, EPA/541/1-89/002, available at: http://www.epa.gov/superfund/programs/risk/ragsa/index.htm

2. U.S. Environmental Protection Agency. 1992. Guidance for Data Useability in Risk Assessment (Part A). Office of Emergency and Remedial Response, Washington, DC. Publication 92857-09A, PB92-93356, available at: http://www.epa.gov/oerrpage/superfund/programs/risk/datause/parta.htm.